STALKMarket

Brand Image Analysis

FINAL DOCUMENTATION

PRI 2019-2020

Fall 2019 batch

Team Stalk Marketers

Amrita Tiwari
Aravind Kumar
Brenda Flores
Gaurav Nepal
Vishal Kalola

30th June 2020

Table of Contents

Acknowledgement

We the Stalk Marketers would like to extend our support and gratitude to all the people who have extended their support for the project with their greatest interest to make it a success.

We would also like to thank the International administration of EPITA for giving us this wonderful opportunity to work in an international environment among different cultures on the Project Research and Innovation (PRI) which makes a strong bond and understanding of the people and their capabilities to flourish in the project.

We had the greatest pleasure to work with Mr. Bill Manos who is our mentor and sponsor for this project, who kept a big trust on us for finishing up the project. He also kept a constant touch on us with the meeting not to check on how the project works but to check on the morale of the team and its efficiency of completing the project. His motivation encouraged us to not divert out of the topic during the tough times the humanity faced in the recent time.

We also would like to thank Professor Olivier Berthet for his support and guidance on the till the competition of our project.

Lastly, we are thankful to each and every one of the team members who worked on this project with such enthusiasm and the level of motivation we had to complete the project on time. We had passed through a tough time which will be engraved in our memory with the support of each other from the team and the rest of the members who has helped us in completing of the project.

Team Stalk Marketers

Abstract

The project Stalk Market was a drop out from the startup company which they taught it could not be completed due to the complexity involved in it. But our team Stalk Market taught of it in a different way that it could be a challenge to our minds. After many brainstorming sessions we came up with a plan which resulted successful working prototype model. It was not an easy task to accomplish the humongous work which we put in the project.

Stalk Market is a tool that processes social media images of the specific brand using image processing technologies. It provides the clients with the insights of how their product is consumed by their customers. The project uses AI and Data technologies where we collect the images of the specific brand and process them to provide the consumer insights by analyzing the elements inside the images, hashtags, captions and the location of the picture taken thus, the brands use the reports from the tool for optimizing their marketing strategies.

It may not be the first tool in the market for process of social media image but we would say it is an efficient one which will generate the desired output according to the customer needs with the statistical analysis with the fully functional Stalk Market tool.

Meet the Stalk Marketers

Stalk Marketers comprised of people from two different countries doing their master's in Software Engineering, Data Science and Computer Security along with our exchange student from Mexico. We had the pleasure with different culture and gained more insights about how they work and think.

Amrita Tiwari

Post: PM & Technical Lead
Profile: MSc. in Computer Science Specializing in Software Engineering @Epita.
Working with international teams and leading them is part of her everyday job. Passionate about designing and developing UI/UX & web applications, also spends quality time in learning new technologies and languages.

Aravind Kumar

Post: Cloud Engineer
Profile: MSc. in Computer Science Specializing in Computer Security and Networks @Epita.
Experienced IT professional of 3 years, effectively implementing new technologies for user-friendly customer experience. Also, curious about knowing the security incidents in computer security world.

Vishal Kalola

Post: Lead Developer
Profile: MSc. in Computer Science Specializing in Software Engineering @Epita.
Software Engineer with 4 years of experience in development and all levels of testing with strong problem solving and analytical skills. Likes learning new development frameworks.

Gaurav Nepal

Post: Data Engineer
Profile: MSc. in Computer Science Specializing in Data Science @Epita. Detail-oriented person who excels at maintaining large databases and analysis by creating detailed reports. Possess great analytical, multitasking skills, and significant ability to work in a team.

Brenda Flores

Post: Researcher
Bachelor Exchange Program @Epita.
Interests in software development and project management. Someone passionate about research and programming. Possess analytical and problem-solving skills.

Vocabulary and Abbreviations

The below table contains the acronyms used in the document

Terms	Meaning
Stalk Market	Tool used for image processing
Client	Clients who uses the tool/ generates the report
B2B	Business to Business
API	Application Program Interface, which is a medium that allows two applications to talk to each other.
Image Processing	Image processing is the use of a digital computer to process digital images through an algorithm
Google Vision	Derive insights from your images in the cloud or at the edge with AutoML Vision or use pre-trained Vision API models to detect emotion, understand text, and more.
CEO	A chief executive officer (CEO) is the highest-ranking executive in a company

Background Information

The general idea of this project is to create a report for the companies where their product is widely used like the location (inside/outside the house, shops, malls, etc.)

Project Objective

Companies now a days tends to see where and who uses their product the most. Let it be from the children to the youngster to the adults on how frequently they use the product and for what purpose. The companies keep a tab on the social media and other platforms where the people post about their product and how they are using it.

The goal of this project is to collect the images posted in the social media like Facebook, Instagram, Twitter, Tumblr etc. using the hashtag and posted in the public group and create a report using the neural networks and Google Vision API to segregate the location, people and other objects in the background and create a detailed analysis of them and give the report to the client who can then market their product according to the data they received to increase their sale of the product.

Business Model

Since it is a B2B application, the cost is based on the subscription type the client chooses and in future after full development of the project and new features the cost of the subscription changes accordingly.

Product Overview

The Problem

The companies now a days are feeling difficult to market their product to attract the customers with the data they possess. The image processing tools available on market focus on the elements of the images and various other details but lacks to give a definite recognition on the overall view of the brand and how their brands are published in image format in the public image banks and social media sites.

The Solution

Social media is one of the most powerful platforms in today's date. People sharing about their daily life, activities, and the products they use is influencing every other user in the sites like Instagram, Facebook, etc. What if we use the same features and collaborate them with image processing technologies to come up with better marketing strategies? What if the brands know how their consumers are sharing the pictures about the products? The answer to all this question leads to one solution i.e. *The Stalk Market tool*. The tool provides the data in the graphical representation of different categories of the elements present in the image like cosmetic etc. The solution will be to

- Exact data of where the brand product is used widely
- Real-time result of the product
- Simulation of how the report is generated with the image analysis technique

Business Case

The main reason for this project is to give an accurate result of the customer product and where it is used widely. The reasons are due to

- Failure to recognise the brand products.
- Even with the results, the brand products cannot conclude on how and where it is used.

The proposed product stalk market will have the following benefits

- Recognises the brand logo of the product.
- The results from the stalk market will be accurate and the reports can be used to boost their product in the market.

Software & Technologies

Software, libraries, and third-party API's to be used

Google Vision AI
Derive insights from your images in the cloud or at the edge with AutoML Vision or use pre-trained Vision API models to detect emotion, understand text, and more.

PyCharm
PyCharm is an integrated development environment used in computer programming, specifically for the Python language.

MySQL and MySQL workbench
MySQL is an open-source relational database management system.
MySQL Workbench is a unified visual tool for database architects, developers, and DBAs.

Photoshop
Adobe Photoshop is a raster graphics editor

Python
Python is an interpreted, high-level, general-purpose programming language.

Django
Django is a Python-based free and open-source web framework that follows the model- template-view architectural pattern.

HTML
Hypertext Markup Language is the standard markup language for documents designed to be displayed in a web browser.

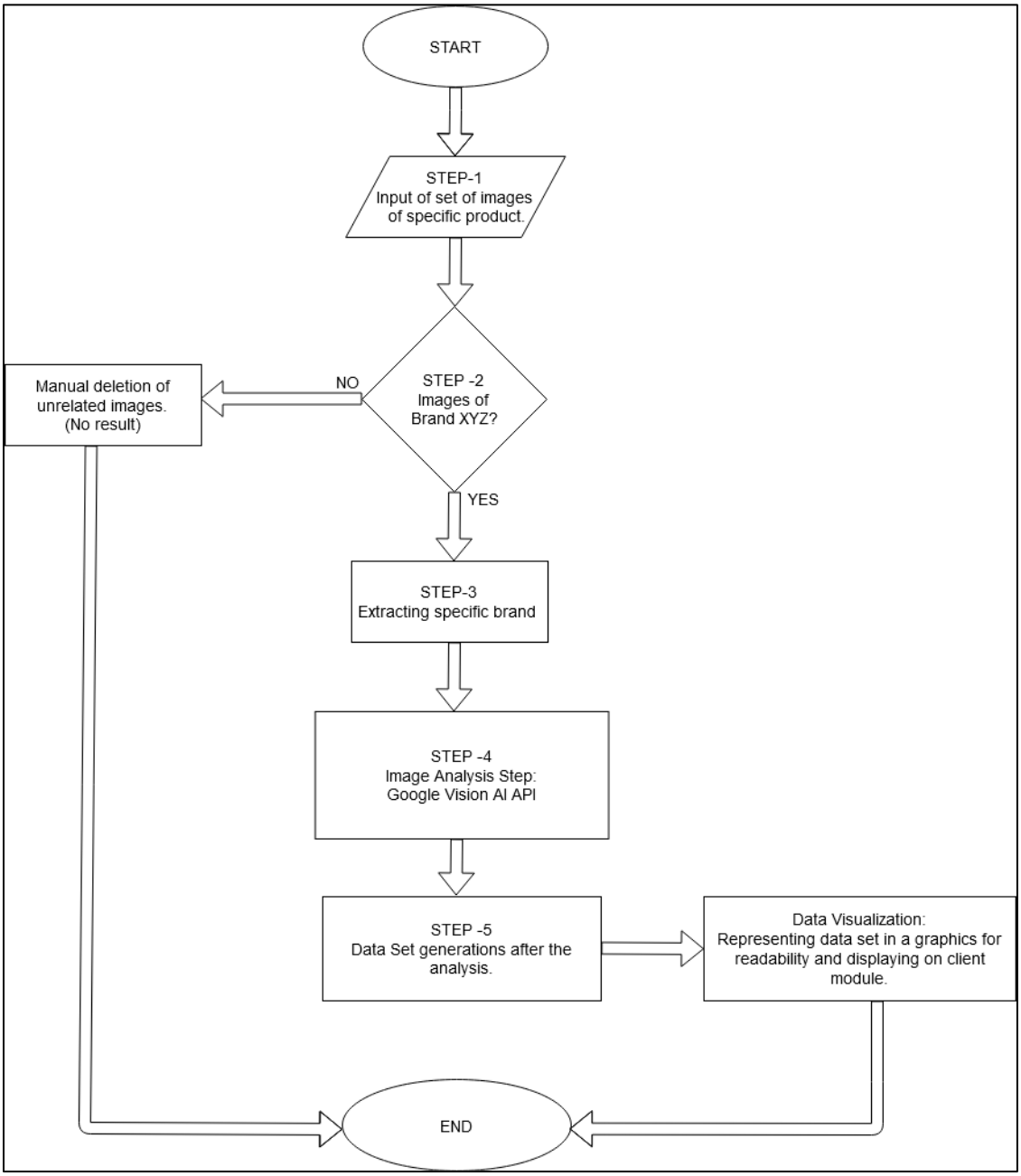
CSS
Cascading Style Sheets is a style sheet language used for describing the presentation of a document written in a markup language like HTML.

JavaScript
JavaScript, often abbreviated as JS, is a programming language that conforms to the ECMAScript specification.

GitHub
GitHub provides hosting for software development and version control using Git

PRODUCT DETAIL

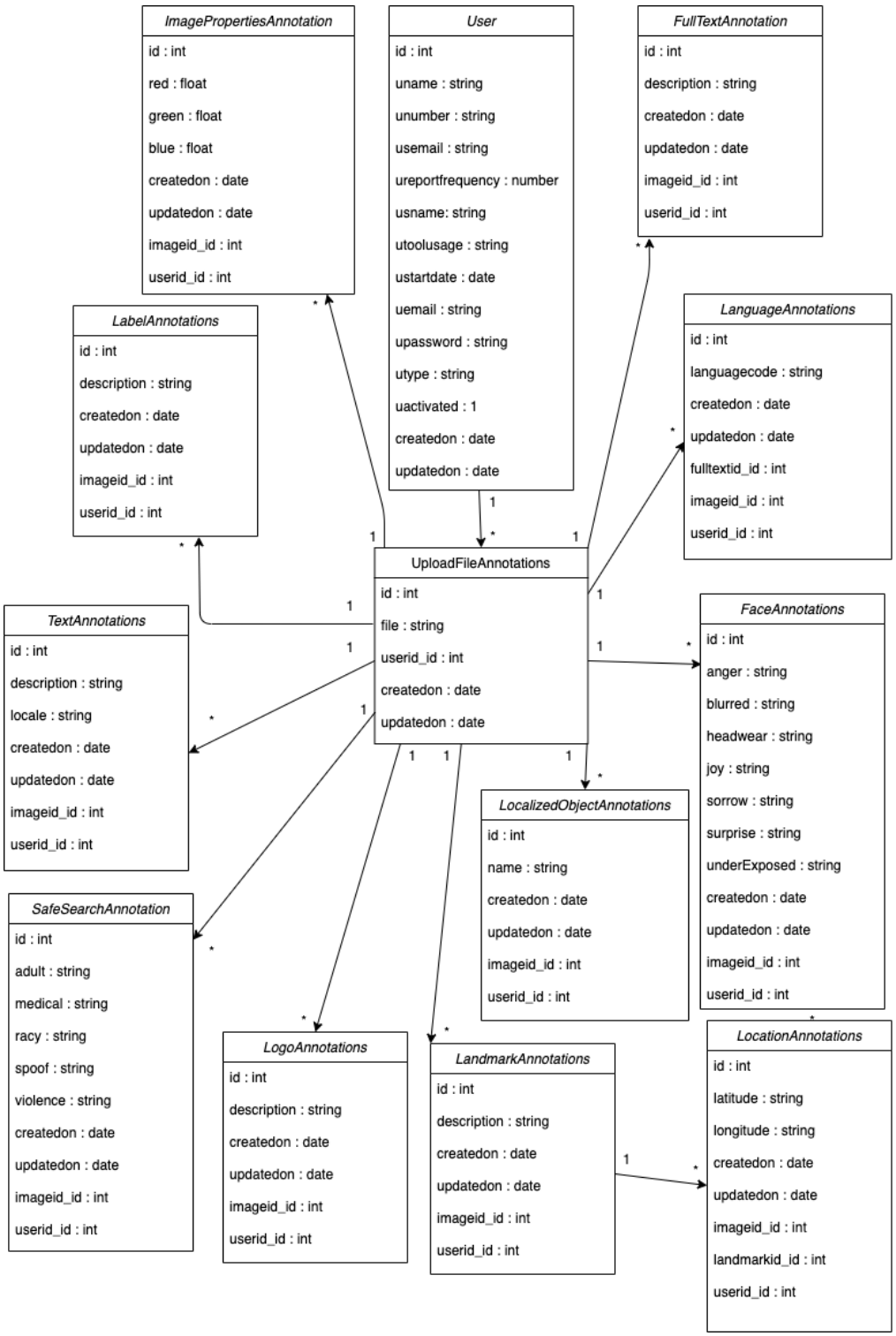
Flowchart of Product



The goal of the tool is to analyze the images and generate the dataset providing the clients with different categorized data. The categories would be related to the elements of the images, the texts used, human emotions detection, etc. All this dataset is represented in graphics which can be used by the clients to make their marketing strategies.

Database Structure

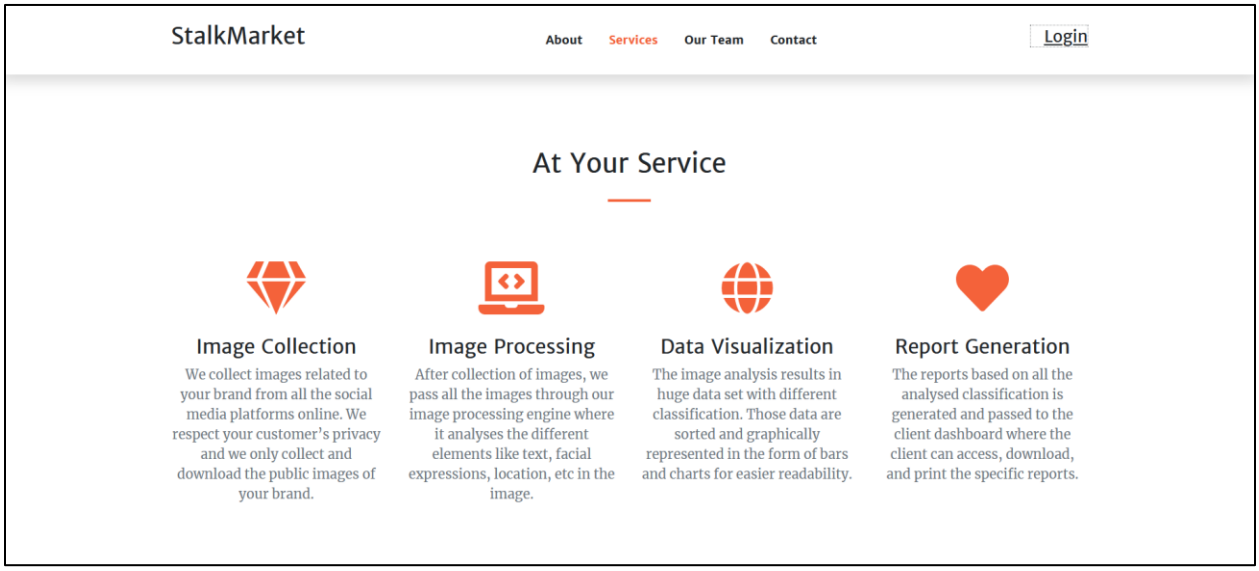
Different tables are created for users, images and for the datasets created after the analysis by categories.

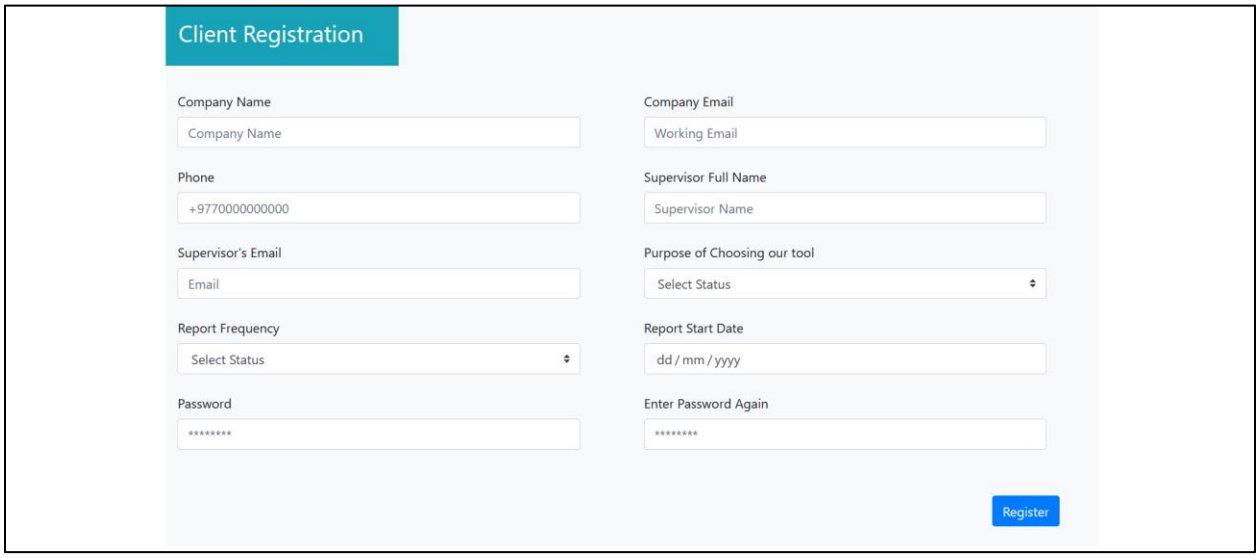


Frontend (Web Interface and UI)

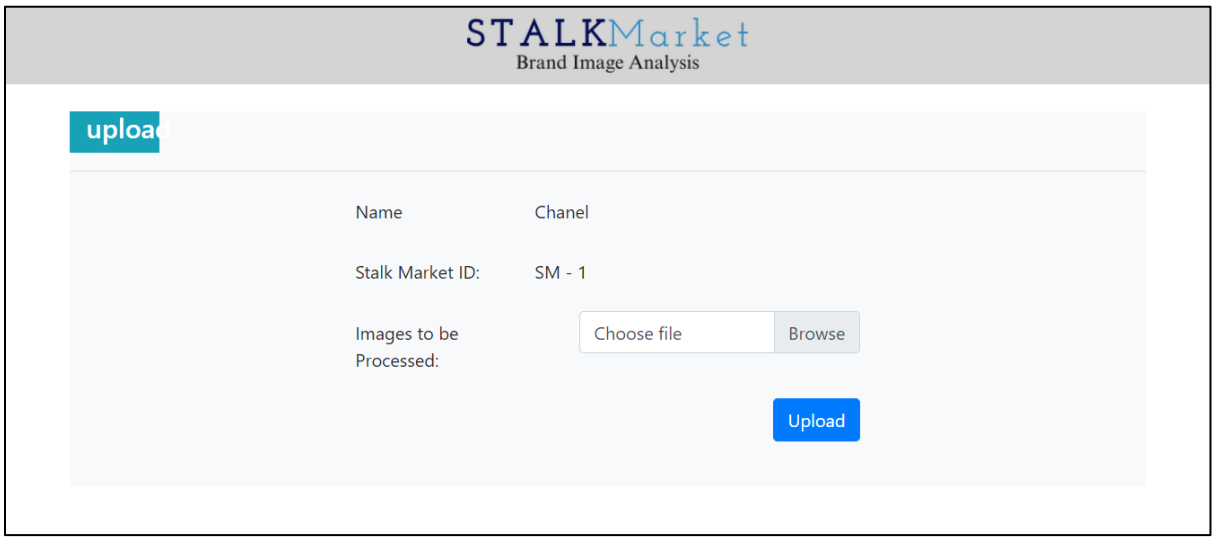
We designed the web site for the project where we display the details of the product and give login options to the admin and the clients. Client registration option is given to new subscribers where they enter their details.

Home Page

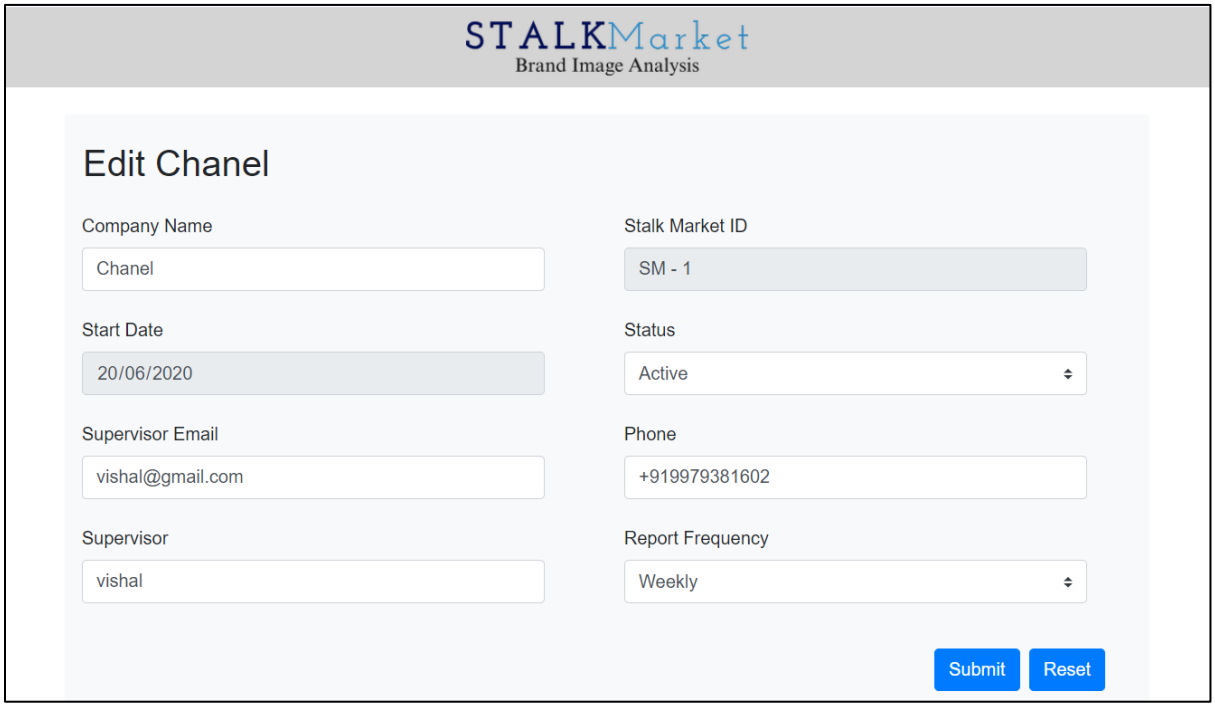




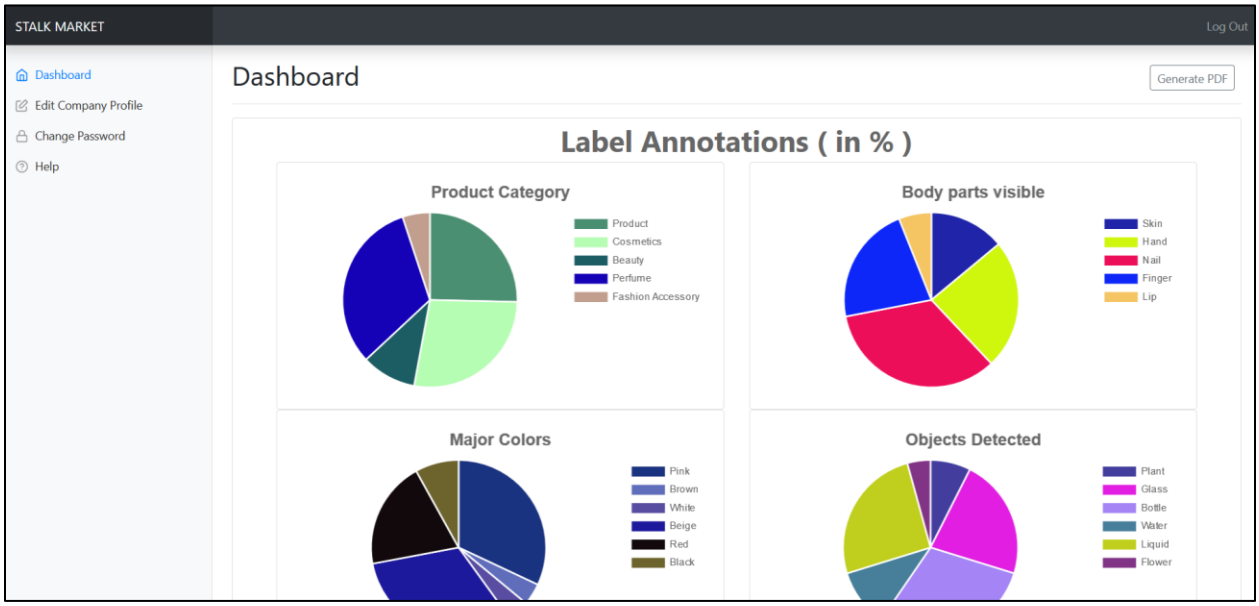
The admin CMS has the dashboard displays the companies that are registered where admins can upload the collected images and run the “Image Processing Operation”. After the process is completed, it generates the dataset and it is updated in the database.

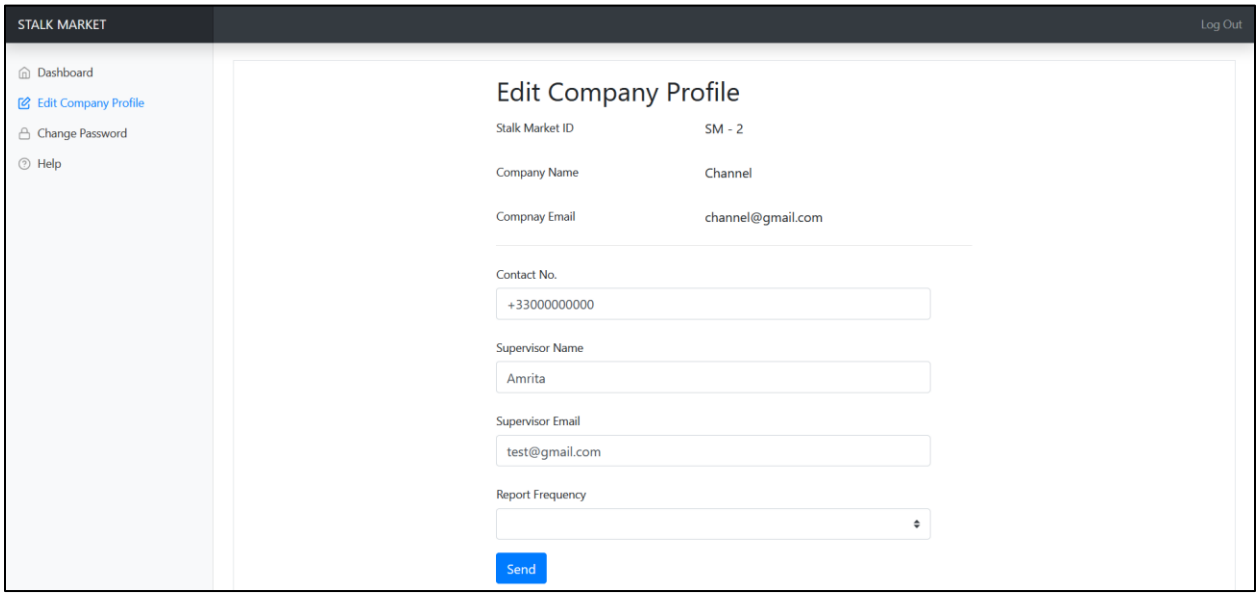



STALKMarket
Brand Image Analysis
Chanel Report Dashboard
CompanyChanel Name
Start Date June 20, 2020
Email vishalkalola196@gmail.c
Supervisorishal
Supervisorishal@gmail.com Email
Stalk Market ID SM-1
Status Activate
Phone +919979381602
Report Frequency Weekly
No. of Report Generated 0


STALKMarket
Brand Image Analysis
Edit Chanel
Company Name
Chanel
Stalk Market ID
SM - 1
Start Date
20/06/2020
Status
Active
Supervisor Email
vishal@gmail.com
Phone
+919979381602
Supervisor
vishal
Report Frequency
Weekly
Submit
Reset

On client side, the dashboard has the recent reports created using the dataset from the database. They can access the monthly reports selecting on the month and can print and download them in pdf formats. They have option to change the password and contact the company for any help.

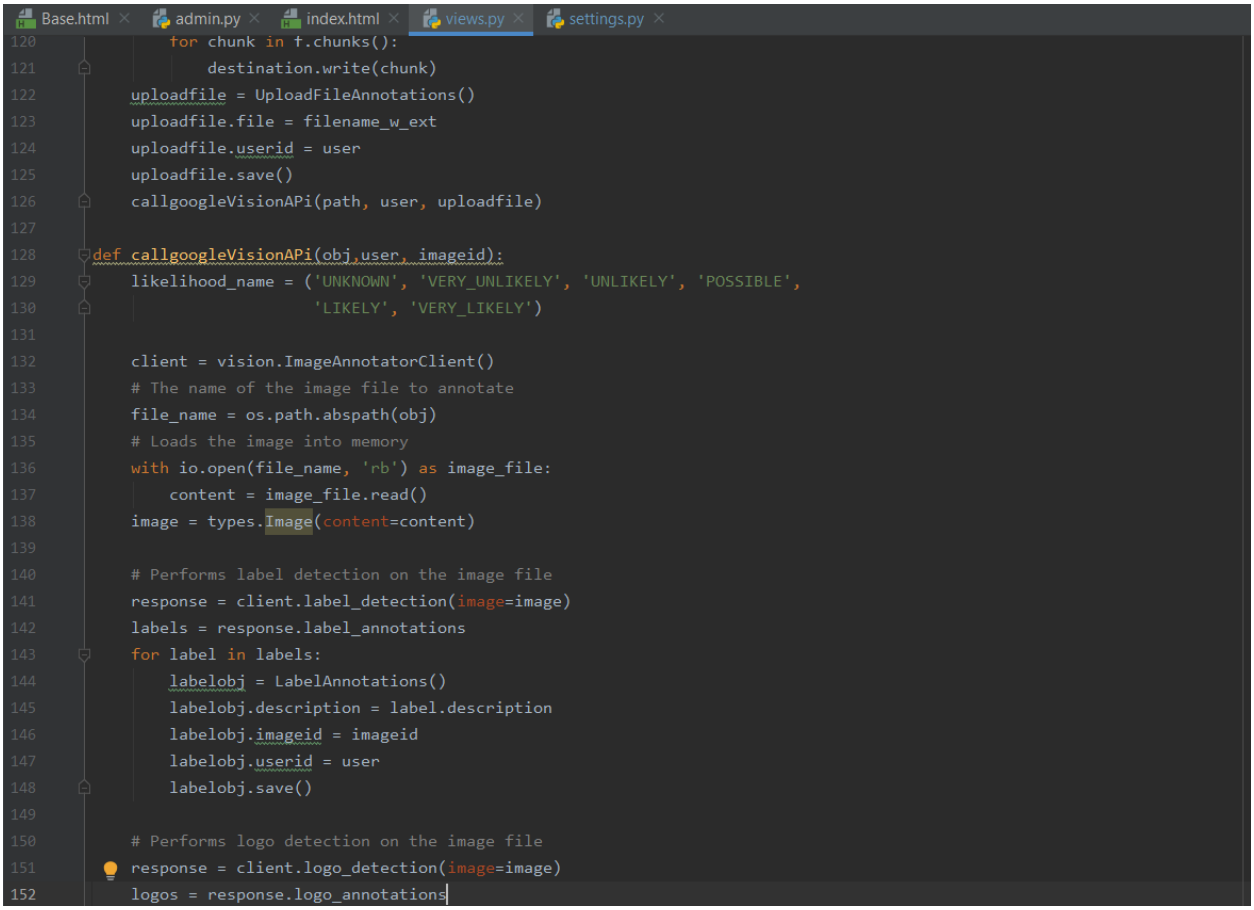




Backend (API Integration)

Backend server developed using Python and Google vision AI API runs the main process of the project i.e. the image processing. As of now, we are using the free version of API which can process 1K images per month which can be upgraded in future as per the requirement. When we upload the images, the images are stored and analyzed that results in dataset production. After the dataset is produced, we delete the images uploaded to free up the memory. The dataset is then used for data visualization.

Image: code snippet of integration of Google Vision API AI API



System Workflow:

Administrator

1. Edit company details
2. Delete the client company
3. Upload the images collected of the specific company
4. Run the image processing operation

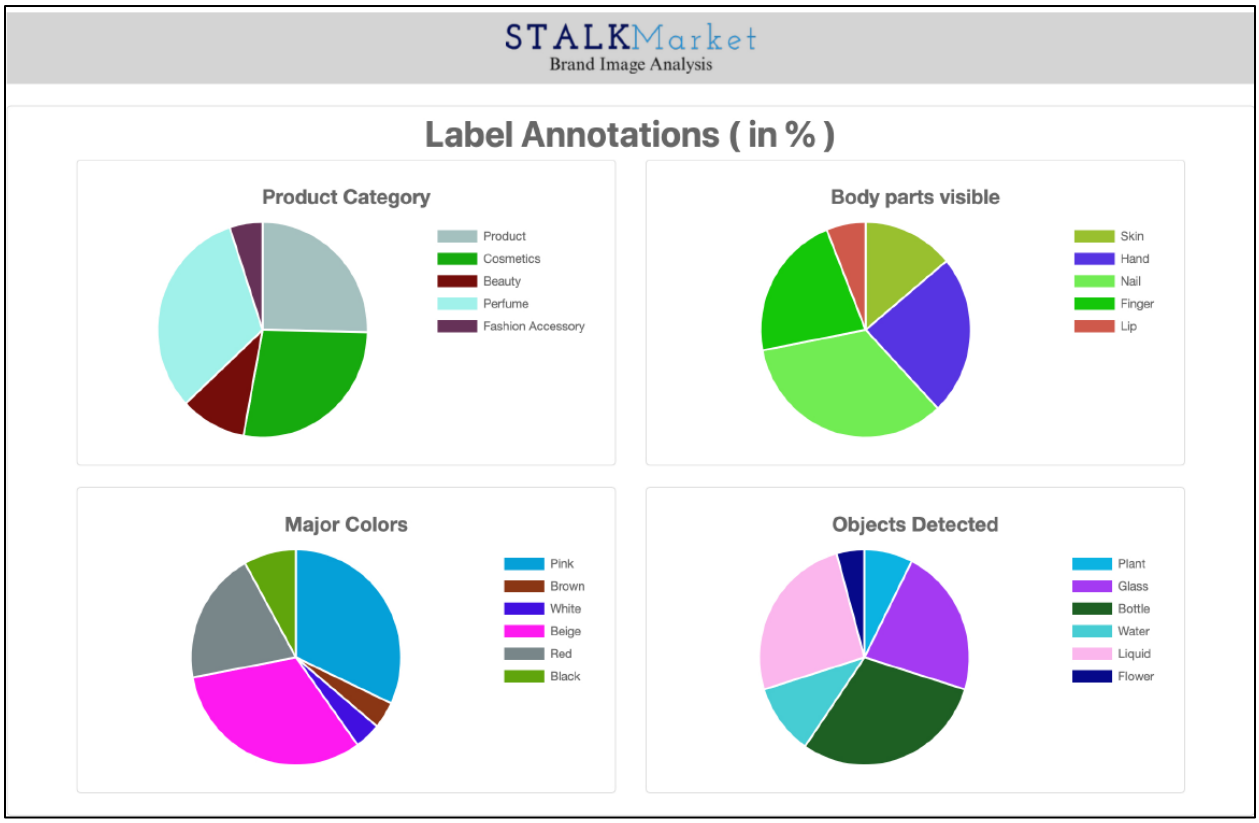
Client

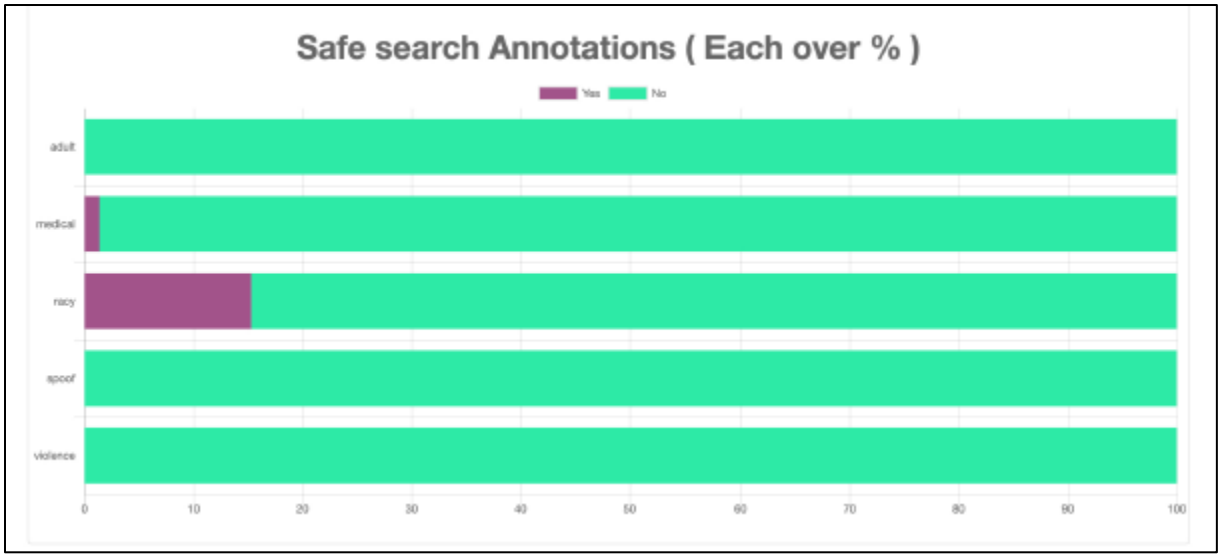
1. Create/Register account
2. Edit the profile
3. Change password
4. View and download reports
5. Message for help

Data Produced (Reports)

The dataset produced is saved in the MySQL tables by categories. But representing those huge amounts of data into user readable format is very important. We adopted data visualization in python to build graphs for each table and display the dataset for the clients where they don't have to spend too much time and energy to understand the data that has been generated.

Part of the report in the pdf form is as below:





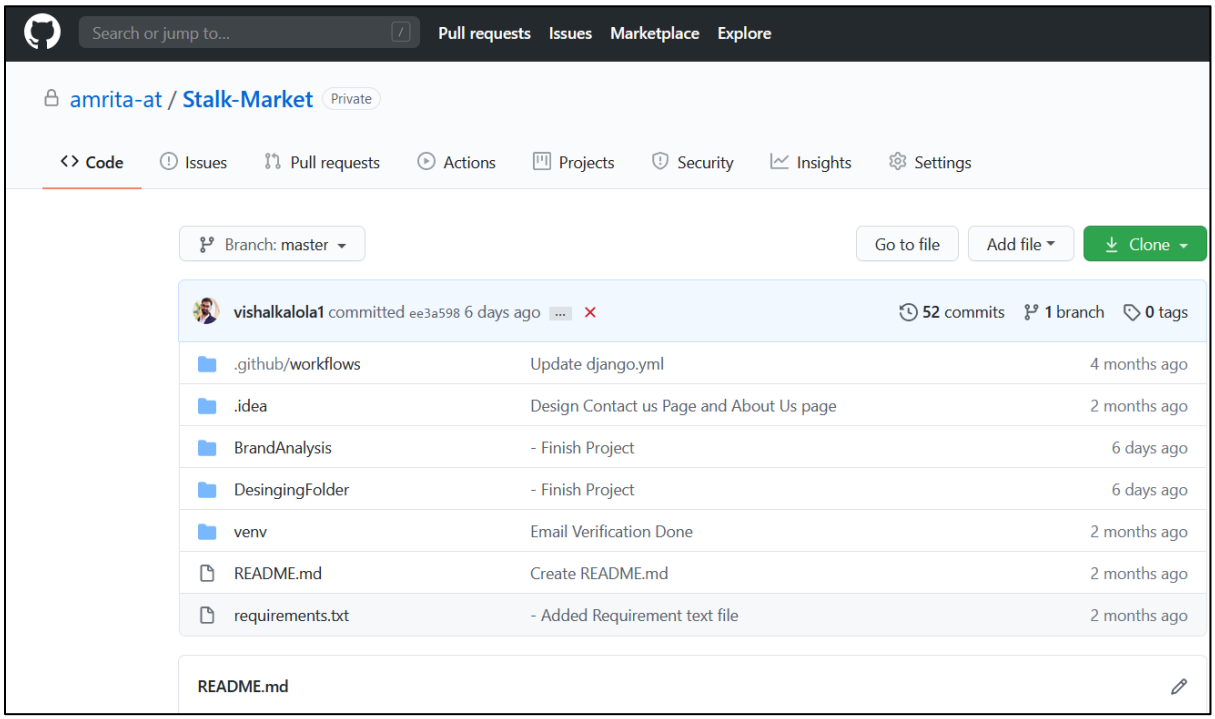
Data Visualization

Code snippet of data visualization in python:

```
        # chart - 1
        labeldata = makeLabelData(user)
        context["labeldata"] = labeldata[0]
        context["labellabels"] = labeldata[1]

        # chart - 2
        piedata = makePieChart(user)
        context["piedata"] = piedata[0]
        context["pielabels"] = piedata[1]

        # Chart - 3
        groupbarchart = makeGroupChart(user)
        context["groupbardata"] = groupbarchart[0]
        context["groupbarlabel"] = groupbarchart[1]
        context["grouplikelyhood"] = groupbarchart[2]

        # Chart - 4
        horizontalbarchart = makeHorizontalGroupChart(user)
        context["horizontalbarchartdata"] = horizontalbarchart[0]
        context["horizontalbarchartlabel"] = horizontalbarchart[1]

        #chart - 5

        data = makelogoAnnotationPie(user)
        context["logodata"] = data[0]
        context["logolabel"] = data[1]
```

Repository

The source code and all the related files are stored in private repository in GitHub with access to every team member.



GDPR

The GDPR sets out 7 main principles to process the personal data in EU. All the data's which we get from the client are hashed and there will no link to a specific person or company. The only private data we collect is during client registration such as company name, contact details which would be saved in encrypted format with us. When the client decides to end the subscription for the tool, the details collected previously would be saved no longer than 2 years.

User Scenario

Scenario 1

A perfume manufacture company CEO wants to see where his product is used by the customers mostly. So, he checks on many stores and gets feedback but was not satisfied with the output which he received. So now he tries the Stalk Market tool to see on the how the product of his company is used. To his surprise the tool gives a statistical analysis of who uses the product (male, female, kids, youngster, etc.) , the place where it is being used (like the restroom, bedroom, etc.) and see these details in the form of graphs and he understands how it is being used. This helped him on to focus on where to sell his product and to whom.

Scenario 2

A computer manufacturing company wants to know where their laptop is being sold the most and to whom. They are using the Stalk Market tool to see the output they expected. After a detailed analysis of the images taken from the social media they understood that the IT companies uses their product the most since it is reliable (From Feedback posted) and then the students who are moving to college or pursuing their studies buys the max. This result helps them to give a quality offer to the students and the company (for the bulk purchase) to sell their product the most.

Scenario 3

It is the time of summer and an ice cream company want to know where to sell their product i.e. the location and which flavor. They use the Stalk Market tool to know the locations of their product where it sells the most and the flavor. And after a detailed study of the images and the background information posted by the people using the company name as a hashtag, they got the result what they expected for and that the season sales

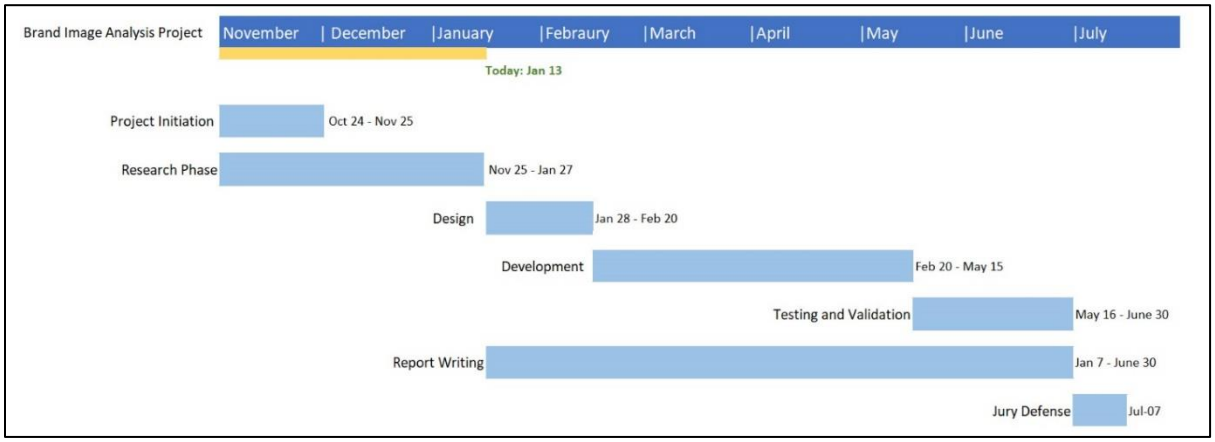
Project Cycle

Project started in November 26th, 2019. Following are the phases of project:

- Project Initiation
- Research Phase
- Design
- Development
- Testing and Validation
- Report Writing
- Jury Defence

Gannt Chart

The milestones and dates of the projects are represented in the Gantt chart as follows:



Risk Management

Technical Complexity

	Risk	Mitigation
1	Bad quality of image will affect the output of the result	Bad quality of image will be negated when using the neural networks.
2	In case of any major upgrade in the tool, loss of data or the mechanism of the tool can be affected.	If major upgrade is happening it will be happening in phases with certain intervals.
3	Neural network implementation is not done since the process takes longer time that expected.	We used API for the image and logo detection and now we are going with manual segregation.

Development tools and hosting of website

	Risk	Mitigation
1	We used the free subscription of the Google Vision AI to analyze the and get the output. After 6 months the free subscription ends.	When it the project goes to live the free subscription will be renewed to paid one.
2	We are hosting the tool in our local host and not hosting live since there is no budget for it.	In future if we have any sponsor to develop the application further it will be hosted online.

Automation of the report

	Risk	Mitigation
1	The automation of the report cannot be done since it is a prototype.	In future if we have any sponsor to develop the application further it will be hosted online.

Future Work and Conclusion

Future Work

Currently, our project has the feature of processing images and generating reports on monthly basis. Following would be the features to be updated in future:

- Optimizing the image upload and generation of reports.
- Adding the subscription types on weekly basis.
- Giving clients the options to choose the number of images to be processed.
- Conducting surveys on what data, companies focus on and expects us to include in reports.
- Including “marketing hints” in the reports, like examples of marketing strategies relating to the dataset.

Conclusion

Marketing is one of the major components in the business regardless of millions of categories. Building effective strategies takes an account of multiple subjects. Image processing can help in marketing the brands in wider vision and get details in deeper level. Social media being very powerful, analyzing images from those platforms would help brands to think being on their customer’s end.

Hence, Stalk Market tool helps in finding all the possible data from the images related to brand, to provide the brands with more detailed view of the images being shared online. Creating effective strategies would be easier with use of our tool. We are delivering the working prototype of the project, but to be used in real companies, the tool must be updated and optimized to the higher level of usage.

Advancement in technology cannot be predicted, it gets updated every moment we blink. Image processing and data visualization is one of the highly trending technologies in current time. Hence combing the leading technologies and the marketing world can give unbelievable results in branding and in other sectors of the industries. Adopting technologies and updating the industrial management is very important if brands must survive in the market.